DOCUMENT RESUME

ED 171 780 TM 009 241

AUTHOR McConnell, John W.; Bowers, Norman D.
TITLE A Comparison of High-Inference and Low-Inference Measures of Teacher Behaviors as Predictors of Pupil Attitudes and Achievements.
PUB DATE Apr 79
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Algebra; *Classroom Observation Techniques; High School Students; Rating Scales; Secondary School Mathematics; Senior High Schools; *Student Attitudes; Student Teacher Relationship; *Teacher Behavior; *Teacher Role

ABSTRACT

Data derived from a year-long study of 43 algebra classes in 13 high schools were collected to assess the differences between high-inference and low-inference measures of teacher behavior as predictors of pupil achievement and attitude change. The high-inference systems rated characteristics such as enthusiasm and clarity while low-inference measures were often frequency counts of objective, specific teaching acts. Student attitudes and achievement were measured at the beginning and end of the year. Achievement measures involved computation, comprehension, and analysis. Attitude measures included liking mathematics versus non-mathematics; mathematics fun versus dull; mathematics easy versus hard; and actual mathematics self concept. Results showed: (1) relationships of low- and high-inference measures to attitude and achievement measures were low; (2) low inference measures produced more significant correlation coefficients with achievement measures than did high-inference measures; (3) high-inference measures produced more significant correlation coefficients with attitude measures than did low-inference measures; and (4) correlations among low-inference measures and achievement measures were greater in magnitude than the correlations among high-inference measures and achievement measures. It was concluded that both types of observation instruments should be used in process-product research. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED171780

A COMPARISON OF HIGH-INFERENCE AND LOW-INFERENCE MEASURES OF TEACHER BEHAVIORS AS PREDICTORS OF PUPIL ATTITUDES AND ACHIEVEMENTS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John McConnell

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

John W. McConnell

Glenbrook South High School
Glenview, Illinois

Norman D. Bowers

Northwestern University
Evanston, Illinois

Paper Presented at the
Annual Meeting of the American Educational Research Association
San Francisco
April 8-12, 1979

TM 009 241

A Comparison of High-Inference and
Low-Inference Measures of Teacher Behaviors as Predictors
Of Pupil Attitudes and Achievements

ABSTRACT

Reviews of process-product research have suggested that both high-inference and low-inference measures should be used to study teacher effects. This study compared selected high-inference and low-inference measures, using data from 43 secondary school algebra classrooms. Three high-inference rating measures and three low-inference observation systems produced the process variables. Achievement and attitude measures were administered at the beginning and end of a school year; residual gains for each classroom were derived as the product measures. Results substantiate the notion that use of both high-inference and low-inference measures can improve the study of relationships among teacher behaviors and pupil outcomes. Surprisingly, low-inference measures seemed to offer somewhat better results than high-inference measures, particularly with achievement criteria. To optimize results in future studies, both types of measures should be used.

A Comparison of High-Inference and
Low-Inference Measures of Teacher Behaviors as Predictors
Of Pupil Attitudes and Achievements

INTRODUCTION

In their reviews of process-product studies, Rosenshine and Furst (1971, 1973) recommended that both high-inference and low-inference measures be used to study teacher effects on pupil learning. This paper presents data derived from a process-product study which integrated both types of teacher behavior measures.

Gage (1969) introduced the concept of observer inference as a means of distinguishing observation systems. Rosenshine (1970a) expanded on the concept by naming observation systems as rating systems or category systems. Low-inference systems require observers to make few interpretive decisions or judgements as they code teacher behavior. The observer determines only which category describes the behavior observed during a small unit of time. Usually a rating of teacher behavior is obtained after a relatively long (20 to 40 minute) time unit. Each system has both advantages and disadvantages.

High-inference systems measure characteristics which can be communicated to the public, and to teachers, as important for good teaching (enthusiasm, clarity, warmth, etc.). These characteristics can be assessed with little expense. Observer training is less extensive than that required by low-inference systems. It has been demonstrated that students can be used to make reliable ratings of teacher behavior (Anderson, Walberg, Welch, 1969). Rating scales

have well-known disadvantages. They are subject to halo effects and to errors due to the lack of a calibration standard. High-inference measures do not translate into specific teacher behaviors. If a construct such as "enthusiasm" is correlated with pupil learning, little direction is provided by the high-inference ratings as to the specific behaviors necessary for a teacher to display the trait.

Low-inference measures of teacher behaviors may provide some solutions to these problems. Since the low-inference measures tend to be frequency counts of objective, specific teaching acts (such as "teacher rebukes student"), they can be translated into programs and prescriptions for teacher education. Although teachers sometimes ridicule the specificity of category systems, that same specificity makes results easier for teachers to accept. Having a low number of teacher-pupil interchanges is less threatening to a teacher than being rated low on "clarity." Work by Flanders (1970) and by Medley and Mitzel (1963) indicated that category systems have the additional virtue of providing frequency counts which can be arithmetically manipulated and factor analyzed into composite measures. These composites can then be constructed to represent global constructs such as "enthusiasm" and "clarity."

Low-inference measures do have disadvantages. First, the problem of defining critical, atomic teacher behaviors has not been solved. The plethora of category systems which have been developed (Simon and Boyer, 1970) testifies to the difficulty of identifying basic units of teacher behavior. Second, category systems are expensive to use. Observers must be trained and their observations monitored to prevent drift from category definitions. Many observations of a classroom are required to discern trends and to record instances of rare, but important and perhaps critical, teacher behaviors.

Rosenshine and Furst have recommended that researchers use both kinds of measures to capitalize on their respective strengths. Further, they indicated that their review of the literature on process-product studies provided some evidence that predictive validity for high-inference measures was higher for high-inference measures than for low-inference measures (1971, p.56). Were this true, a researcher could use the high-inference measures to identify those broader constructs of teacher performance which might relate to pupil change. The low-inference measures in the same study would provide the concrete, specific teacher behaviors which could be considered as defining the broader constructs. This paper considers data from a year-long study of algebra classrooms and concentrates on the differences between high-inference and low-inference measures of teacher behavior as predictors of pupil learning and attitude change.

PROCEDURE[1]

Data on pupil achievement and attitude and teacher behavior were collected from forty-three algebra classes in thirteen Chicago suburban high schools during one school year. Algebra classes had been selected because both course objectives and the content of algebra texts used in the schools were relatively standard from class to class and school to school. The attitude and achievement items that formed the basis for the product measures were derived from the National Longitudinal Study of Mathematical Abilities (NLSMA) of the School Mathematics Study Group (Romberg and Wilson, 1969). These items provided multidimensional measures of learning and attitude. The items appeared to be more satisfactory than those from item pools available to researchers in a number of other subject areas.

---

[1]More detailed information on these aspects of the study and on the high-inference measures may be found in McConnell (1977, 1978b).

Product Measures. Student attitudes and achievements were measured at the beginning and end of the year. The spring scores on three achievement scales and four attitude scales were regressed against fall scores to remove variability due to initial class differences. The residuals provided the product measures for further analyses. The use of class means as the unit of analysis and multiple predictor variables for each criterion measure provided unusually pure measures of pupil learning and attitude change (McConnell, 1978b). The product measures considered in this study are listed below. Descriptions are adapted from Romberg and Wilson (1969).

Achievement:

Computation - Items designed to require straight-forward manipulation of problem elements according to rules the subjects presumably have learned. Emphasis is on performing operations, not upon deciding which operations are appropriate.

Comprehension - Items designed to require either recall of concept and generalizations or transformation of problem elements from one mode to another. Emphasis is upon demonstrating understanding of concepts and their relationships.

Analysis - Items designed to require a nonroutine application of concepts.

Attitude:

Math vs. Non-Math - This scale is designed to measure how well a student likes mathematics and considers it important in relation to other subjects.

Math Fun vs. Dull - This scale is designed to measure the pleasure or boredom a student experiences with regard to mathematics, both in the absolute sense and comparatively with other subjects.

Math Easy vs. Hard - This scale is designed to measure the ease or difficulty which a student associates with mathematics.

Actual Math Self-Concept - This scale is designed to measure how a student sees himself in relation to mathematics.

High-Inference Process Measures. Three of the high-inference measures of teacher behavior were based on constructs synthesized by Rosenshine and Furst. These were the Classroom Process Questionnaire (CPQ), the Process Rating Sheet (PRS), and the Teacher Characteristics Sheet (TCS). The CPQ required pupils to rate their teacher on items measuring clarity, task-orientation, variability, enthusiasm, opportunity to learn, indirectness, criticism, structuring, types of questions, probing, and difficulty. The PRS was used by an observer to rate the teacher on the same constructs. The observer also used the TCS, which contained items from the Teacher Characteristics Study (Ryans, 1960), and measured teacher warmth, enthusiasm, task-orientation, and concreteness.

In summary, the high-inference variables used in this study are on the following list. For the remainder of this paper, identifiers for these high-inference variables will be in capitals.

Classroom Process Questionnaire

CPCLART	Clarity
CPQUEST	Questioning
CPVART	Variability
CPENTHT	Enthusiasm
CPDIFFT	Difficulty
CPCRITT	Criticism
CPUSET	Use of Student Ideas
CPSTRUCT	Structuring
CPOPPT	Opportunity to Learn
CPTASKT	Task-Orientation

Process Rating Sheet

PR1	Clarity
PR2	Variability
PR3	Enthusiasm
PR4	Businesslike
PR5	Uses Student Ideas
PR6	Criticism
PR7	Structures Lesson
PR8	Higher Cognitive Questions
PR9	Probing Questions
PR10	Difficulty of Lessons

Teacher Characteristics Sheet

TCSX	Kindly, understanding, friendly
TCSY	Responsible, systematic, businesslike
TCSZ	Stimulating, original, enthusiastic
TCSC	Concrete, practical

Low-Inference Process Measures. Three low-inference observation instruments were used during each visit to an algebra classroom. They were chosen to span the affective and cognitive activities encountered in day-by-day mathematics instruction. The three category systems were OScAR4V, Wright-Proctor, and the Pupil Involvement Inventory.

The OScAR4V focused on two major categories of verbal interaction: the statement and interchange (Simon and Boyer, 1970). Teacher statements were classified as affective, substantive, or procedural. Interchanges, particularly those relating to content, were classified as to entry and exit. An entry to an interchange could be a pupil question, a teacher question getting a student to elaborate, a teacher question intending pupil thought and divergent response, or a teacher question aimed at a specific response. The exits were based on teacher responses to the student portion of the interchange: acceptance, disapproval, or neutral. Fourteen variables were formed from frequencies on the OScAR4V. These are listed below. For the remainder of the paper, the low-inference measures will be given in lower-case. This will distinguish the low-inference measures from the high-inference measures, which are indicated in capitals, as discussed above.

ostot	Total statements and interchanges

Teacher Statements:

ospst	Problem-structuring
osdes	Describing
osinf	Informing
oscons	Considering
osreb	Rebuking
osdir	Directing
ostea	Total teacher statements

Interchange Entrances:

osdiv	Divergent teacher questions
oselab	Elaborating teacher questions
osconv	Convergent teacher questions

Interchange Exits:

ossa	Teacher response to pupil - support and approval
oscr	Criticism and rejection
osnon	Nonsubstantive interchanges

The work of Wright (1959) and Wright and Proctor (1961) was used to structure the two other observation systems. Wright and Proctor had set up an observational system which classified verbal behaviors in mathematics classrooms along three dimensions: content, process, and attitude. The three dimensions proved to be unwieldy in the tryouts of instruments prior to the start of this study. Therefore, the content and process dimensions were separated from the affective. The resulting content/process instrument will be termed the Wright-Proctor throughout this paper, and the attitude-oriented instrument will be called the Pupil Involvement Inventory.

The content dimension of the Wright-Proctor referred specifically to the mathematics under discussion. There were three major categories: Fundamentals, Relations, and Applications. The first two categories differentiated the body of mathematical knowledge at the command of the pupils (Fundamentals) from the development of new relations (Relations). The Applications category picked up statements dealing with the use and place of mathematical systems in specific problems and in a historical context.

The process dimension of the Wright-Proctor also had three major divisions. Syllogistic focused on inference, particularly with respect to Analyzing and Synthesizing (high cognitive) behaviors. Classificatory categories included Specialization and Generalization. (In the sense of their definitions, these

were considered lower cognitive.) The third category, Relevant, was used when mathematical information was presented, but which was not connected to an apparent logical sequence.

The classification of verbal behaviors did require judgement on the part of the observer. By limiting the instrument to two dimensions, it was possible to record four tallies each minute, twice the pace reported by Wright (1959). The class was observed for ten seconds. The observer used five seconds to classify the primary activity during the ten second observation. The judgements required during the observation period were not easy to make, even though this instrument was considered low-inference: the judgements required that the observer be trained in mathematics and be familiar with the sequence of the algebra curriculum.

Eight variables were derived from the Wright-Proctor for this study. They are listed below.

wphico	Fundamentals - high cognitive
wploco	Fundamentals - low cognitive
wpnrt	New relations
wpapp	Applications
wpttlk	Proportion of teacher talk
wpwork	Work time
wpansy	Analysis and synthesis
wpsyn	Synthesis

The affective component of the original Wright-Proctor was substantially revised for the Minnesota National Laboratory Five State System (Wright, 1967). The revision merged Wright and Proctor's affective dimension with Flanders Interaction Analysis. Categories were defined in such a way that distinctions unique to mathematics teaching could be made. Nevertheless, it was possible to equate the Pupil Involvement Categories with the Flanders Interaction Analysis to exploit derived variables, such as indirect/direct ratio, commonly used in interaction research with FIA (Flanders, 1970). One tally was made every five

seconds. The variables used in this study included

piex	Teacher clarifies, encourages, summarizes
picont	T contacts, checks
piconf	T confronts, seeks
pichall	T challenges, jolts
piinfo	T informs, lectures
pidirct	T directs
piid	Indirect/direct ratio

Observation Schedule. A trained observer visited each class twice during each semester. During the class period, five minutes were allowed for initial class organization. The observer then conducted six five-minute observation sessions with three instruments. A typical session included:

5 minutes	Wait
5 minutes	Wright-Proctor
5 minutes	Pupil Involvement Inventory
5 minutes	OScAR4V
5 minutes	Wright Proctor
5 minutes	Pupil Involvement Inventory
5 minutes	OScAR4V
Remainder of period	High-Inference Rating Scales (PRS, TCS)

The order of three low-inference instruments was permuted at each visit. Complete observation data were obtained on 45 classes; complete pupil data were obtained for 43 classes.

Analyses. All 53 process measures were correlated with the seven product measures. Significant correlations ($P < .05$) were examined to determine whether the 24 high-inference measures were more predictive of pupil change than the 29 low-inference.

RESULTS

The significant correlations ($P < .05$) between product measures and process measures are given in Table 1 and Table 2. Table 1 gives the correlations of pupil learning with teacher behavior measures. Table 2 shows the relationships among pupil attitude change and teacher behavior measures.

The low-inference measures appeared more predictive of pupil learning than did the high-inference measures (Table 1). For each of the three product measures, variables from the catagory systems accounted for more criterion variance than did the variables developed from the rating systems. Further, more of the low-inference measures correlated with product measures (16 significant correlation coefficients of a possible 87 correlation coefficients, or 18%) than did high-inference measures (8 significant of 72 correlation coefficients, or 11%).

Results for the attitude product measures show few differences between low-inference measures and high-inference measures in predictive validity (Table 2). Low-inference measures provided correlations with the attitude product measures which were as substantial as those provided by high-inference measures. More significant correlation coefficients ($P < .05$) were obtained among attitude measures and high-inference measures (20 significant correlation coefficients of 96 correlation coefficients, or 21%) than among attitude product measures and low-inference measures (20 significant of 116 correlation coefficients, or 17%).

Table 1

SIGNIFICANT CORPELATIONS (P<.05) OF LOW-INFERENCE AND HIGH-INFERENCE MEASURES OF TEACHER BEHAVIOR WITH STUDENT LEARNING MEASURES IN 43 NINTH GRADE ALGEBRA CLASSROOMS[a]

Absolute Value of Correlation	Achievement Measure		
	Computation	Comprehension	Analysis
.45 - 1.00		wpsyn (.46)	
.40 - .44			wpsyn (.44) osdiv (.40)
.35 - .39			
.30 - .34	osnon (-.34) osdiv (.34) osreb (-.32) ospst (.32) wpapp (.31) CPTASKT (.31)	wphico (.34) wpansy (.33) osnon (-.32)	wpansy (.30)
.25 - .29	CPQUEST (.28) PR4 (.28) TCSY (.26)	wpttlk (.29) ospst (.29) PR1 (.28) CPTASKT (.28) CPCLART (.26)	pidrct (.28) PR9 (.27) pidirct (-.26) CPENTRT (.26)

[a]High-inference measures are indicated in CAPITALS; low-inference measures are indicated in lower-case.

Table 2

SIGNIFICANT CORRELATIONS (P<.05) OF LOW-INFERENCE AND HIGH-INFERENCE MEASURES OF TEACHER BEHAVIOR WITH SELECTED MEASURES OF CLASS ATTITUDE CHANGE IN 43 NINTH GRADE ALGEBRA CLASSROOMS[a]

Absolute Value of Correlation	Attitude Measure			
	Math Fun vs. Dull	Math vs. Nonmath	Math Easy vs. Hard	Actual Math Self-Concept
.45 - 1.00				CPTASKT (.46)
.40 - .44	wpnrt (.40)		piconf (.41)	
.35 - .39				CPCLART (.38)
	PR1 (.36)	osdiv (.36) piid (.36)		
	piid (.35)	CPENTHT (.35) PR4 (.35)		CPENTHT (.35)
.30 - .34	osdes (.34)		wpnrt (.34)	
				osdes (.33) TCSZ (.32)
	CPENTHT (.31) TCSZ (.30) piconf (.30)	TCSZ (.31) CPCLART (.30) wphico (-.30) oscr (.30)		wpnrt (.31) PR1 (.30)
.25 - .29		piinf (-.29)	osdes (.29) piid (.29)	
	CPOPPT (.28) CPCLART (.28) PR8 (.28)	osinf (-.28)	pichall (.28)	
			TCSZ (.27) PR1 (.27)	CPOPPT (.27)
	osdiv (.26) CPTASKT (.26) wphico (-.26)	osdes (.26) PR8 (.26)		

[a]High-inference measures are indicated in CAPITALS; low-inference measures are indicated in lower-case.

DISCUSSION

Results suggest the following:

1. Relationships of both low- and high-inference measures to attitude and achievement measures were low (.25-.40). They were little different from correlations reported in many other studies over the last twenty years.

2. The low-inference measures produced more significant correlation coefficients ($P < .05$) with achievement product measures than did the high-inference measures. This was reversed for the attitude product measures.

3. Correlations among low-inference measures of teacher behavior and achievement product measures were greater in magnitude than the correlations among high-inference measures and achievement product measures. This was not true for the attitude product measures.

These conclusions run counter to Rosenshine and Furst's conclusion from their review of the literature that the high-inference measures of teacher behavior would be more predictive of pupil change. Further work (not reported here) indicates that relating high-inference measures to low-inference measures in the same study does not yield a clear atomization of critical teacher behaviors.

Nevertheless, both types of teacher observation instruments should be used in process-product research. Previous papers dealing with this study of algebra classroom focused on the relationships among high-inference measures of teacher behaviors and pupil outcomes (McConnell, 1977) and teacher attitudes (McConnell, 1978a). The results of those papers and the results detailed here indicate that the techniques which were incorporated into this study to assist in the identification of a limited number of critical teacher behaviors have instead exposed the vital complexity of the classroom. Different teacher behaviors have different effects

on the multiple outcomes desired for pupil attitudes and achievements. Pupils and observers, ostensibly measuring the same constructs, have different talents for discerning the influences of different teacher attitudes. Low-inference measures, rather than clarifying and defining the global constructs represented by the high-inference measures, may be representing a distinct set of important teacher behaviors. Further process-product studies should continue to use multiple sources of observation and multiple measures of teacher process to examine teacher effects on pupil learning and attitude change.

REFERENCES

Anderson, Gary J., Walberg, Herbert J., and Welch, Wayne W. "Curriculum Effects on the Social Climate of Learning: A New Representation of Discriminant Functions." American Educational Research Journal vi, May, 1969, 315-328.

Campbell, Donald T. and Fiske, Ronald W. "Convergent and Divergent Validation by the Multitrait-Multimethod Matrix." Psychological Bulletin. March, 1959, 81-105.

Flanders, Ned A. Analyzing Teacher Behavior. Reading, Massachusetts: Addison-Wesley Publishing Company, 1970.

Gage, N. L. "Teaching Methods," in Encyclopedia of Educational Research, Fourth Edition, Robert L. Ebel (ed.). American Educational Research Association, 1969, 1446-1458.

McConnell, John W. "Relationships Between Selected Teacher Behaviors and Attitudes/Achievements of Algebra Classes." Paper presented at the Annual Meeting of the American Educational Association, New York, 1977. (ERIC ED 141-118).

McConnell, John W. "Relations Between Teacher Attitudes and Teacher Behavior in Ninth-Grade Algebra Classes." Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, 1978(a) (ERIC ED 152-751).

McConnell, John W. Relationships Among High-Influence Measures of Teacher Behaviors and Student Achievement and Attitude in Ninth Grade Algebra Classes. Unpublished Ph.D. dissertation, Northwestern University, 1978(b).

Medley, D. M. and Mitzel, H. E. "Measuring Classroom Behavior by Systematic Observation" in Handbook of Research on Teaching. N. L. Gage (ed.) Chicago: Rand McNally, 1963, 247-328.

Medley, D. M. and Mitzel, H. E. "The Scientific Study of Teacher Behavior" in Theory and Research on Teaching. Arno A. Bellack, ed. New York: Teachers College Press, Columbia University, 1963, 79-90.

Romberg, Thomas and Wilson, James W. The Development of Tests. National Longitudinal Study of Mathematical Abilities, Report #7. James W. Wilson, Leonard S. Cahen, Edward G. Begle, eds. Palo Alto: Stanford University, 1969.

Rosenshine, Barak. (1970a). "Evaluation of Classroom Instruction." Review of Educational Research 40, April, 1970, 279-300.

Rosenshine, Barak. (1970b). "The Stability of Teacher Effects Upon Student Achievement." Review of Educational Research 40, December, 1970, 647-662.

Rosenshine, Barak and Furst, Norma. "Research in Teacher Performance Criteria" in Research in Teacher Education. B. Othanel Smith (ed.) Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1971, 37-72.

Rosenshine, Barak and Furst, Norma. "The use of Direct Observation to Study Teaching" in Second Handbook of Research on Teaching. Robert M. W. Travers (ed.) Chicago: Rand McNally & Company, 1973, 122-183.

Ryans, D. G. Characteristics of Teachers. Washington, D. C.: American Council on Education, 1960.

Simon, Anita and Boyer, E. Gil. Mirrors for Behavior II: An Anthology of Classroom Observation Instruments. Special Edition of Classroom Interaction Newsletter. Philadelphia: Research for Better Schools, Inc., 1970.

Wright, Elisabeth Muriel. "Development of an Instrument for Studying Verbal Behaviors in a Secondary School Mathematics Classroom." Journal of Experimental Education XXVIII, December, 1959, 103-121.

Wright, Elisabeth Muriel J. "A Rationale for Direct Observation of Verbal Behaviors in the Mathematics Classroom," in Psychological Problems and Research Methods in Mathematics Training, Rosalind L. Feierabend and Philip DuBois (eds.), 38-46.

Wright, Elisabeth Muriel. Teacher-Pupil Interaction in the Mathematics Classroom. Minnesota National Laboratory, Minnesota State Department of Education, Technical Report No. 76-5, 1967.

Wright, Elisabeth Muriel and Proctor, Virginia H. Systematic Observation of Verbal Interaction as a Method of Comparing Mathematics Lessons. U.S. Office of Education, Cooperative Research Project No. 816. St. Louis: Washington University, 1961.